JOURNAL OF BACTERIOLOGY, May 1995, p. 2953–2955 Vol. 177, No. 10
0021-9193/95/$04.00+0

The "Universal" Leucine Codon CTG in the Secreted Aspartyl Proteinase 1 (*SAP1*) Gene of *Candida albicans* Encodes a Serine In Vivo

THEODORE C. WHITE,[1] LORI E. ANDREWS,[2] DAVID MALTBY,[2] AND NINA AGABIAN[1]*

Intercampus Program in Molecular Parasitology, Department of Stomatology and the Oral AIDS Center, University of California—San Francisco and Berkeley, San Francisco, California 94143-1204,[1] and Mass Spectrometry Research Facility, University of California at San Francisco, San Francisco, California 94143-0446[2]

Received 23 January 1995/Accepted 16 March 1995

A number of *Candida* species possess a $tRNA^{Ser}$-like species that recognizes CTG codons that normally specify leucine (Leu) in the universal code of codon usage. Mass spectrometry and Edman sequencing of peptides from the secreted aspartyl proteinase isoenzyme (Sap1) demonstrate that positions specified by the CTG codon contain a nonmodified serine (Ser) in *Candida albicans*.

Deviations from the universal code have been identified in a variety of nuclear and mitochondrial genomes in the last 15 years (reviewed in reference 12). Recent analysis of codons from several *Candida* species has demonstrated that one of the universal Leu codons, CTG, is associated with unusual translational events. Unorthodox tRNAs that recognize the CTG codon and that are charged in vitro with Ser have been identified from both the pathogenic fungus *Candida albicans* (14) and the distantly related yeasts *C. cylindracea* (20) and *C. maltosa* (15, 21). These tRNAs show extensive homology to $tRNA^{Ser}$ species from these and other yeasts yet contain a CAG anticodon, while a $tRNA^{Leu}$ that recognizes the CTG codon has not been identified (15, 16). The CTG Ser codon has been identified, by in vitro translation and tRNA sequencing, in 6 of 14 yeast species, including *C. albicans*, *C. cylindracea*, *C. melibiosica*, *C. parapsilosis*, *C. rugosa*, and *C. zeylanoides* (11). Recently, the introduction of a $tRNA^{Leu}$ recognizing the CTG codon was found to be toxic in *C. albicans* (6).

The amino acids incorporated at CTG codons have not been extensively characterized. In *C. cylindracea*, sequence analysis of a cDNA for secreted lipase 1 revealed 19 CTG codons which appear to specify Ser, as determined by partial amino acid sequencing of the isolated protein (5). In *C. albicans*, the incorporated amino acid has not been identified, although the addition of the unusual $tRNA^{Ser}$ from *C. albicans* to in vitro translation reaction mixtures results in proteins with slower than expected mobilities on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (13, 14, 17). While multiple Ser residues can decrease the relative mobility of proteins (2), it is unlikely that the replacement of one or two Leu residues with Ser during in vitro translations would result in the mobility differences seen, suggesting that a modified or nonstandard amino acid may be incorporated by the CTG $tRNA^{Ser}$ in *C. albicans*. Alternatively, this $tRNA^{Ser}$ may have an unusual effect on translational initiation or elongation in the in vitro reactions (14). The low frequency of CTG codons in genes from *C. albicans* (one or two per gene) in comparison with the frequency in genes from *C. cylindracea* (19 per gene) would be consistent with the incorporation of a modified or nonstandard amino acid in *C. albicans* proteins.

During our analysis of the *SAP* gene family, we noticed that *SAP1* and *SAP7* contain CTG codons (Table 1). While conditions which induce expression of the *SAP7* gene have not been detected (3), the *SAP1* gene is transcribed in the O cell type of strain WO-1, as determined by Northern (RNA) blot analysis of total RNA (7, 9, 10). Both SDS-PAGE and isoelectric focusing of culture supernatants from O cells can be used to resolve at least three different Sap isoenzymes, including Sap1 (18). The differences in electrophoretic mobility of the three isoenzymes cannot be explained by the primary sequence (18), suggesting that differences may be due to the incorporation of nonstandard or posttranslationally modified amino acids. Sequence analysis of both of the alleles of the *SAP1* gene from strain WO-1 shows that both alleles include the CTG codon, and S1 analysis shows that both alleles are expressed (7).

To identify the amino acid specified by CTG in *C. albicans*, we isolated large quantities of Sap1 from *C. albicans* for peptide analysis. Undiluted culture supernatant from O cells grown in proteinase-inducing medium (YCB-BSA) contains Sap1, -2, and -3 (18). The supernatant was isoelectrically focused in a Rotofor (as previously described) (18), and fractions shown to contain the Sap1 protein by immunoblotting were electrophoresed through preparative SDS–8% PAGE. The Sap1 protein band was excised from the gel, electroeluted, lyophilized, suspended in buffer, and acetone precipitated. This purified Sap1 protein was digested to completion with endoproteinase Asp-N (3.2 μg/μl, enzyme/substrate ratio = 1:18.25; Boehringer Mannheim) in 100 mM NH_4HCO_3, pH 7.8, at 37°C for 18 h. The resulting peptides were separated and mass analyzed by liquid chromatography-mass spectrometry (LCMS). The C_{18} LC eluate was run through a UV detector (ABI model 783A), and subsequently the flow was split, with 10 μl going to the mass spectrometer (Fisons Platform; Fisons, Manchester, United Kingdom), equipped with an electrospray source, and the rest being collected for protein sequencing. The data were acquired and processed by using MassLynx software. Sequence analysis of peptide fractions was performed on a gas phase sequencer (ABI model 470A) equipped with an on-line PTH analyzer (ABI model 130A) (1).

The Asp-N digestion of the Sap1 protein results in a peptide, positions 235 to 243 of the mature protein, that contains the

* Corresponding author. Mailing address: Intercampus Program in Molecular Parasitology, University of California San Francisco, 3333 California St., Suite 150, San Francisco, CA 94118. Phone: (415) 476-6845. Fax: (415) 476-0664. Electronic mail address: agabian@itsa.ucsf.edu.

TABLE 1. Leu codon usage in the *SAP* genes of *C. albicans*

Gene[a]	Frequency of Leu codon:						Total
	CTG	CTA	CTC	CTT	TTA	TTG	
SAP1	1		2	9	11	9	32
SAP2				8	16	9	33
SAP3				9	13	9	31
SAP4			2	8	14	8	32
SAP5			3	8	13	8	32
SAP6			3	8	13	9	33
SAP7	3		2	6	19	13	43
Total	4	0	12	56	99	65	236

[a] References for gene sequences are as follows: *SAP1*, reference 4; *SAP2*, 19; *SAP3*, 18; *SAP4*, 7; *SAP5*, *SAP6*, and *SAP7*, 8.

CTG-encoded amino acid at position 243 (4). The UV trace and total ion current trace of the peptides are shown in Fig. 1A. The average molecular mass of the CTG amino acid-containing peptide is expected to be 1,034.2 Da if it contains Leu and 1,008.1 Da if it contains Ser. The LCMS trace does contain a peptide fraction (fraction 21 retention time, 51 min) with a measured average molecular mass of 1,008.1 Da (Fig. 1B) and does not contain a fraction with a mass of 1,034.2 Da,

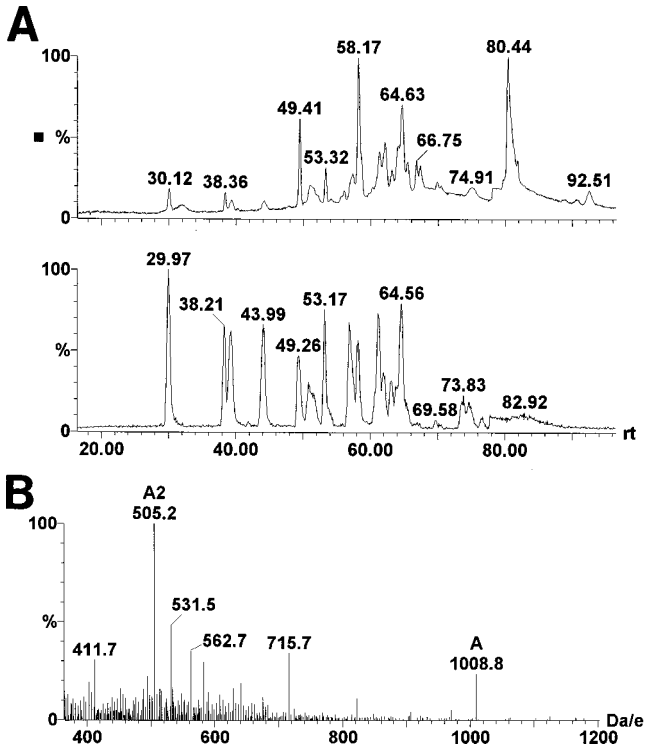

FIG. 1. (A) The top panel shows a UV trace at 215 nm of LCMS of Sap1 protein digested with the endoproteinase Asp-N. The *x* axis represents the retention time (rt) in minutes. The *y* axis represents the percent relative intensity compared with the intensity of the most intense peak. Numbers above peaks indicate peak retention times. The bottom panel shows an MS total ion current trace. The *x* axis represents the retention time in minutes, adjusted for the lag between UV scanner and MS. The *y* axis represents the percent relative intensity compared with the intensity of the most intense peak. Numbers above peaks indicate peak retention times. (B) Spectrum of fraction 21 showing the doubly (A2) and singly (A) charged ions of the 1,008.1-Da peptide. The *x* axis represents the mass-to-charge ratio (*m*/*z*) expressed as daltons/e (e = change of an electron). The *y* axis represents the percent relative intensity compared with the intensity of the most intense peak.

TABLE 2. Expected and observed peptide characteristics

Peptide and parameter	Mol wt[a] (avg)	Peptide sequence[a]
CTG amino acid containing (235–243[b])		
Predicted for Leu	1,034.2	DAFQAELKL
Predicted for Ser	1,008.1	DAFQAELKS
Fraction 21	1,008.1	DAFQAELKS
CTC amino acid containing (196–202[b])		
Predicted for Leu	788.0	ITLNSLK
Predicted for Ser	761.9	ITLNSSK
Peptide mixture	788.2[c]	ITLNSLK....[d]
	1,111.3	YTSASNIAALT
	1,239.5	.....QNLGTP

[a] Predicted molecular weights and sequences were based on the gene sequence. Actual values (fraction 21 and peptide mixture) for molecular weights were determined by MS as the means of two or more determinations. Actual peptide sequences were determined by Edman sequencing. Residues of interest are underlined.

[b] Amino acid coordinates based on predicted sequence of mature protein.

[c] Fraction contained peptides with three different molecular weights.

[d] Peptide sequences contained several different amino acids in each position, as shown.

suggesting that fraction 21 contains the peptide in which position 243 is Ser. The Edman sequence analysis of this fraction (Table 2) confirms that the 1,008.1-Da species is the predicted peptide containing a Ser residue at position 243. The above-described analysis demonstrates that the Ser is not modified in vivo, as a covalent modification would result in a peptide with an altered mass and the modification would be detected in the sequence analysis. A minor amino acid species at position 243 was identified as Leu by Edman sequencing, although its presence is most likely the result of trailing from Leu at position 241.

The *SAP1* gene also includes two CTC codons (Table 1), which normally encode Leu. One of these codons (position 201 of the mature protein) was also analyzed by LCMS and Edman sequencing. Peptides were prepared by double digestion with endoproteinase Asp-N and trypsin. The peptide of interest (positions 196 to 202) is expected to be 788.0 Da if it contains a Leu and 761.9 if it contains a Ser at position 201 (Table 2). A column fraction was identified which contains a mixture of three peptides, 788.2, 1,111.3, and 1,239.5 Da. The identification of a 788.2-Da peptide suggests that a Leu residue is found at position 201. To confirm this identification, the fraction was rechromatographed in a shallower gradient, which reduced the levels of the 1,111.3-Da species. Edman sequencing of the fraction gave a mixed sequence (Table 2), in which the predicted sequence of the peptide at positions 196 to 202, including the Leu at position 201, could be identified. No Ser was identified at this position. A second sequence can be identified that corresponds to the 1,111.3-Da contaminating peptide. Traces of other amino acids could also be identified from the analysis (Table 2). The identification of the 788.2-Da peptide and the amino acid sequencing showing Leu at position 201 demonstrate that this CTC codon encodes Leu.

Our results show that the CTG codon in *C. albicans* is translated in vivo to encode a nonmodified Ser, which is apparently unaltered during secretion of the Sap isoenzyme. Our previous analyses of *SAP1* gene sequences, which identified CTG codons in both alleles, and our S1 analysis, which demonstrated expression of both alleles, show that the Ser residues in the protein are encoded by the CTG codons in the gene (7). The related leucine codon, CTC, encodes leucine in the Sap1 protein as expected. This is the first identification of the amino

acid encoded by the CTG codon and translated in vivo in *C. albicans*. Similar analyses have been performed in very few other *Candida* species known to contain altered CTG codons (5, 15). The presence of nonmodified Ser in these organisms suggests that the CTG might encode Ser in all species in which CTG codons are altered.

The identification of serine residues in peptides from the lipase 1 gene of *C. cylindracea* (5) was the first analysis of proteins translated in vivo from species which contain the unusual CTG $tRNA^{Ser}$. Unlike the current study, in which both alleles of the *SAP1* gene have been sequenced (7), the comparison of the lipase protein and cDNA sequences did not include an analysis of both alleles of the lipase gene (5), raising the possibility that allelic differences might have explained the Leu-Ser differences. The only other in vivo amino acid analysis used DNA transformation of *C. maltosa* to introduce plasmids overexpressing a cloned cytochrome P-450 gene containing a CTG codon (15).

The identification of Ser residues at these protein positions does not explain the unusual mobilities of the Sap isoenzymes (18), especially the mobility of Sap2, which does not contain CTG codons, nor does it explain the slower mobilities of proteins translated in vivo in the presence of these $tRNA^{Ser}$ species which recognize CTG codons.

It is likely that the CTG codon originally encoded a Leu and that in an ancestor of these related *Candida* species a $tRNA^{Leu}$ was replaced with a modified $tRNA^{Ser}$. We have looked for indirect evidence of that substitution by inspecting the location of CTG codons in highly conserved proteins (e.g., calmodulin). In each case that we have examined, the CTG codon was located in a nonconserved region of the protein, and neither Leu nor Ser residues were present at that position in other species (data not shown).

The identification of Ser as the translation product of CTG codons is important for the development of molecular genetics in *C. albicans*. The function of foreign gene sequences appears to be profoundly affected by the presence of CTG codons in the gene, as evidenced by the inability of standard reporter genes from heterologous sources to function in *C. albicans* (17a) and in *C. maltosa* (15, 21). The replacement of nonpolar, hydrophobic Leu amino acids for noncharged, polar Ser residues would most likely disrupt the three-dimensional folding of the protein and may perturb key amino acids necessary for function. Similar alterations will occur when *Candida* genes are expressed in other organisms (21), although the low frequency of CTG codon usage (Table 1) should mitigate those effects.

This work was supported in part by NIH grants 1RO1AI-33317 and PO1-DE-07946 to N.A.; a University of California University Wide AIDS Research Program grant from the state of California, no. K92-SF-011, and NIH grant 1RO1DE11367 to T.C.W.; and NIH NCRR BRTP grant RR01614 to the Mass Spectrometry Facility at University of California at San Francisco under the direction of Alma Burlingame.

REFERENCES

1. **Atherton, D., J. Fernandez, M. DeMott, L. Andrews, and S. M. Mische.** 1993. Techniques in protein chemistry, vol. IV, p. 409–418. Academic Press, New York.
2. **Early, A. E., J. G. Williams, H. E. Meyer, S. B. Por, E. Smith, K. L. Williams, and A. A. Gooley.** 1988. Structural characterization of *Dictyostelium discoideum* prespore-specific gene D19 and of its product, cell surface glycoprotein PsA. Mol. Cell. Biol. **8:**3458–3466.
3. **Hube, B., M. Monod, D. A. Schofield, A. J. P. Brown, and N. A. R. Gow.** 1994. Expression of seven members of the gene family encoding secretory aspartyl proteinases in *Candida albicans*. Mol. Microbiol. **14:**87–99.
4. **Hube, B., C. J. Turver, F. C. Odds, H. Eiffert, G. J. Boulnois, H. Kochel, and R. Ruchel.** 1991. Sequence of the *Candida albicans* gene encoding the secretory aspartate proteinase. J. Med. Vet. Mycol. **29:**129–132.
5. **Kawaguchi, Y., H. Honda, M. J. Taniguchi, and S. Iwasaki.** 1989. The codon CUG is read as serine in an asporogenic yeast *Candida cylindracea*. Nature (London) **341:**164–166.
6. **Leuker, C., and J. Ernst.** 1994. Toxicity of a heterologous leucyl-tRNA (anticodon CAG) in the pathogen *Candida albicans*—in vivo evidence for non standard decoding of CUG codons. Mol. Gen. Genet. **245:**212–217.
7. **Miyasaki, S. H., T. C. White, and N. Agabian.** 1994. A fourth secreted aspartyl proteinase gene (*SAP4*) and a *CARE2* repetitive element are located upstream of the *SAP1* gene in *Candida albicans*. J. Bacteriol. **176:**1702–1710.
8. **Monod, M., G. Togni, B. Hube, and D. Sanglard.** 1994. Multiplicity of genes encoding secreted aspartic proteinases in *Candida* species. Mol. Microbiol. **13:**357–368.
9. **Morrow, B., T. Srikantha, J. Anderson, and D. R. Soll.** 1993. Coordinate regulation of two opaque-phase-specific genes during white-opaque switching in *Candida albicans*. Infect. Immun. **61:**1823–1828.
10. **Morrow, B., T. Srikantha, and D. R. Soll.** 1992. Transcription of the gene for a pepsinogen, PEP1, is regulated by white-opaque switching in *Candida albicans*. Mol. Cell. Biol. **12:**2997–3005.
11. **Ohama, T., T. Suzuki, M. Mori, S. Osawa, T. Ueda, K. Watanabe, and T. Nakase.** 1993. Nonuniversal decoding of the leucine codon CUG in several *Candida* species. Nucleic Acids Res. **21:**4039–4045.
12. **Osawa, S., T. H. Jukes, K. Watanabe, and A. Muto.** 1992. Recent evidence for evolution of the genetic code. Microbiol. Rev. **56:**229–264.
13. **Santos, M., D. R. Colthurst, N. Wills, C. S. McLaughlin, and M. F. Tuite.** 1990. Efficient translation of the UAG termination codon in *Candida* species. Curr. Genet. **17:**487–491.
14. **Santos, M. A., G. Keith, and M. F. Tuite.** 1993. Non-standard translational events in *Candida albicans* mediated by an unusual seryl-tRNA with a 5′-CAG-3′ (leucine) anticodon. EMBO J. **12:**607–616.
15. **Sugiyama, H., M. Ohkuma, Y. Masuda, S. M. Park, A. Ohta, and M. Takagi.** 1995. In vivo evidence for non-universal usage of the codon CUG in *Candida maltosa*. Yeast **11:**43–52.
16. **Suzuki, T., T. Ueda, T. Yokogawa, K. Nishikawa, and K. Watanabe.** 1994. Characterization of serine and leucine tRNAs in an asporogenic yeast *Candida cylindracea* and evolutionary implications of genes for tRNA(Ser)CAG responsible for translation of a non-universal genetic code. Nucleic Acids Res. **22:**115–123.
17. **Tuite, M. F., P. A. Bower, and C. S. McLaughlin.** 1986. A novel suppressor tRNA from the dimorphic fungus *Candida albicans*. Biochim. Biophys. Acta **866:**26–31.

17a. **White, T., and N. Agabian.** Unpublished data.

18. **White, T. C., S. H. Miyasaki, and N. Agabian.** 1993. Three distinct secreted aspartyl proteinases in *Candida albicans*. J. Bacteriol. **175:**6126–6133.
19. **Wright, R. J., A. Carne, A. D. Hieber, I. L. Lamont, G. W. Emerson, and P. A. Sullivan.** 1992. A second gene for a secreted aspartate proteinase in *Candida albicans*. J. Bacteriol. **174:**7848–7853.
20. **Yokogawa, T., T. Suzuki, T. Ueda, M. Mori, T. Ohama, Y. Kuchino, S. Yoshinari, I., Motoki, K. Nishikawa, S. Osawa, et al.** 1992. Serine tRNA complementary to the nonuniversal serine codon CUG in *Candida cylindracea*: evolutionary implications. Proc. Natl. Acad. Sci. USA **89:**7408–7411.
21. **Zimmer, T., and W.-H. Schunck.** 1995. A deviation from the universal genetic code in *Candida maltosa* and consequences for heterologous expression of cytochromes P450 52A4 and 52A5 in *Saccharomyces cerevisiae*. Yeast **11:**33–41.